The Unimodular Intersection Problem

Shmuel Onn * Pauline Sarrabezolles †

Abstract

We show that finding minimally intersecting \( n \) paths from \( s \) to \( t \) in a
directed graph or \( n \) perfect matchings in a bipartite graph can be done in
polynomial time. This holds more generally for unimodular set systems.

Keywords: integer programming, combinatorial optimization, unimodu-
lar, graph, path, matching

1 Introduction

Let \( G \) be a directed graph on \( d \) edges with two distinct vertices \( s \) and \( t \) and let \( n \)
be a positive integer. Given \( n \) paths from \( s \) to \( t \), an edge is critical if it is used by
all paths. Here we study the problem of finding \( n \) paths from \( s \) to \( t \) with minimum
number of critical edges. A variant of the problem, where an edge is called shared if
it is used by more than one path, and the goal is to make a choice with minimum
number of shared edges, was recently studied and shown to be NP-hard in [1].

We briefly mention one motivation for this problem, adapted from [1]. We need to
make a sensitive shipment from \( s \) to \( t \). In the planning stage, \( n \) paths are chosen and
prepared. Then, just prior to the actual shipment, one of these paths is randomly
chosen and used. An adversary, trying to harm the shipment and aware of the
prepared paths but not of the actual path that is finally chosen, will try to harm a
critical edge, used by all paths. Therefore, we protect each critical edge with high
cost, and so our goal is to choose paths with minimum number of critical edges.

Let \( A \) be the vertex-edge incidence matrix of \( G \) and let \( b \) be the vector in the ver-
tex space with \( b_s = -1 \), \( b_t = 1 \), and \( b_v = 0 \) for any other vertex \( v \). For \( x^1, \ldots, x^n \in \mathbb{R}^d \)
let \( x^1 \land \cdots \land x^n \in \mathbb{R}^d \) and \( x^1 \lor \cdots \lor x^n \in \mathbb{R}^d \) be the coordinate-wise minimum and
maximum of \( x^1, \ldots, x^n \). Finally, for \( x \in \mathbb{R}^d \) let \( |x| := \sum_{i=1}^d |x_i| \). Then our problem
becomes the following nonlinear integer programming problem defined by \( A, b, n \):

\[
\min \{|x^1 \land \cdots \land x^n| : x^k \in \mathbb{Z}^d, 0 \leq x^k \leq 1, Ax^k = b, k = 1, \ldots, n\} \tag{1}
\]

*Technion - Israel Institute of Technology, Haifa, Israel. Email: onn@ie.technion.ac.il
†Ecole des Ponts, Paris, and Technion, Haifa. Email: pauline.sarrabezolles@gmail.com
In this article we solve this problem, even in the more general context where $A$ is an arbitrary totally unimodular matrix. We establish the following theorem.

**Theorem 1.1** The intersection problem (1) over any totally unimodular matrix $A$, any integer right-hand side $b$, and any positive integer $n$, is polynomial time solvable.

Another consequence of this theorem is the following. Let $G$ be a bipartite graph on $d$ edges and let $n$ be a positive integer. Given $n$ perfect matchings in $G$, an edge is critical if it is used by all matchings. The problem of finding $n$ perfect matchings with minimum number of critical edges is again a special case of our theorem with $A$ the vertex-edge incidence matrix of $G$ and $b$ the vector in the vertex space with $b_v = 1$ for every vertex $v$. So this problem can also be solved in polynomial time.

Before proceeding with the proof in the next section, we briefly point out two variants of problem (1) where even deciding if the optimal value in the problem is 0 is NP-complete, even for fixed $n = 2$. First, if for each $k$ we have a system $Ax^k = b^k$ with the same matrix $A$ but its own right-hand side $b^k$, then, this is hard even if $A$ is the vertex-edge incidence matrix of a directed graph $G$. Indeed, for $k = 1, 2$, let $s_k$ and $t_k$ be distinct vertices in $G$ and let $b^k$ be the vector forcing $x^k \in \{0, 1\}^d$ satisfying $Ax^k = b^k$ to be a path from $s_k$ to $t_k$. Then the optimal objective function value of problem (1) is 0 if and only if for $k = 1, 2$ there are paths from $s_k$ to $t_k$ which are edge-disjoint, which is a classical NP-complete problem to decide, see for example [2]. Second, if $A$ is not totally unimodular, then this is hard even if $A$ has a single row. Indeed, let $a_1, \ldots, a_d$ be given positive integers with even sum, and consider the NP-complete problem of deciding if there is a partition, that is, $I \subseteq \{1, \ldots, d\}$ with $\sum_{i \in I} a_i = \sum_{i \notin I} a_i$. Let $A := [a_1, \ldots, a_d]$, let $b := \frac{1}{2} \sum_{i=1}^d a_i$, and let $n := 2$. Then, if $I$ provides a partition, then $x^1, x^2$ defined by $x^1_i := 1 - x^2_i := 1$ if $i \in I$ and $x^1_i := 1 - x^2_i := 0$ if $i \notin I$, are feasible in (1) with value 0, and if $x^1, x^2$ are feasible in (1) with value 0, then $I := \text{supp}(x^1) := \{i : x^1_i \neq 0\}$ provides a partition.

## 2 Proof

We now prove Theorem 1.1. First, we check the feasibility of $x \in \mathbb{Z}^d$, $0 \leq x \leq 1$, $Ax = b$. This can be done in polynomial time by linear programming since $A$ is totally unimodular. If this system is infeasible then so is our intersection problem and we are done. So assume it is feasible. Now consider the auxiliary program

$$
\min \left\{ \sum_{k=1}^{n-1} k|y^k| + n^2 d |y^n| : y^k \in \mathbb{Z}^d, 0 \leq y^k \leq 1, k = 1, \ldots, n, \sum_{k=1}^{n} Ay^k = nb \right\}. \quad (2)
$$

We begin by proving two claims.

**Claim 1.** Suppose $x^1, \ldots, x^n$ are feasible in (1), with objective value $q$. Define $y^k := \vee \{x^{i_1} \land \cdots \land x^{i_k} : 1 \leq i_1 < \cdots < i_k \leq n\}$, $k = 1, \ldots, n$. 

so \( y^k \in \{0,1\}^d \) with \( y_j^k = 1 \) if and only if \( x_{i_1}^k = \cdots = x_{i_k}^k = 1 \) for some \( i_1 < \cdots < i_k \). Then \( y^1, \ldots, y^n \) are feasible in (2) with objective value \( p \) satisfying \( p < n^2d(q+1) \).

**Proof.** First, by definition \( y^k \in \{0,1\}^d \) for \( k = 1, \ldots, n \). Second, we claim that \( \sum_{k=1}^n y^k = \sum_{k=1}^n x^k \). By definition, \( y^k \leq \cdots \leq y^1 \). Consider any \( 1 \leq j \leq d \). If \( y_j^1 = 0 \) then \( \sum_{i=1}^n y_j^i = 0 = \sum_{i=1}^n x_j^i \) holds. Suppose then that this is not the case and let \( k \) be the largest index with \( y_j^k = 1 \). Then there are \( i_1 < \cdots < i_k \) such that \( x_{i_1}^1 = \cdots = x_{i_k}^1 = 1 \) but \( x_j^i = 0 \) for all other \( i \), and therefore we have again

\[
\sum_{i=1}^n y_j^i = \sum_{i=1}^k y_j^i = k = \sum_{r=1}^k y_j^r = \sum_{i=1}^n x_j^i .
\]

So \( \sum_{k=1}^n Ay^k = \sum_{k=1}^n Ax^k = nb \). Therefore \( y^1, \ldots, y^n \) are indeed feasible in (2). Third, we note that \( |y^k| = |x^1 \wedge \cdots \wedge x^n| = q \), and therefore, as claimed,

\[
p = \sum_{k=1}^{n-1} k|y^k| + n^2d|y^n| \leq (n-1)^2d + n^2dq < n^2d(q+1) .
\]

**Claim 2.** Suppose \( y^1, \ldots, y^n \) are optimal in (2). Then we can find in polynomial time \( x^1, \ldots, x^n \) satisfying

\[
x^k \in \mathbb{Z}^d, \ y^n \leq x^k \leq y^1, \ \sum_{i=1}^n y^i \leq \sum_{i=1}^k x^i \leq \sum_{i=1}^k y^i , \ Ax^k = b , \ k = 1, \ldots, n .
\]

**Proof.** First, we claim that \( y^n \leq \cdots \leq y^1 \). Suppose indirectly there are \( i \) and \( j \) with \( y_j^{i+1} > y_j^i \). Then \( y_j^i = 0 \) and \( y_j^{i+1} = 1 \). Define \( z^i := y^i + 1_j \), \( z^{i+1} := y^{i+1} - 1_j \), where \( 1_j \) is the \( j \)-th unit vector, and \( z^k := y^k \) for \( k \neq i, i+1 \). We claim that \( z^1, \ldots, z^n \) are feasible in (2) with objective function value smaller than \( y^1, \ldots, y^n \), which is a contradiction: indeed, \( z^k \in \{0,1\}^d \) for all \( k \); further, \( \sum_{k=1}^n Az^k = \sum_{k=1}^n y^k \) so \( \sum_{k=1}^n Az^k = \sum_{k=1}^n Ay^k = nb \) and last, \( \sum_{k=1}^{n+1} k|z^k| + n^2d|z^n| < \sum_{k=1}^{n-1} k|y^k| + n^2d|y^n| \).

We now prove by induction on \( r \) that we can find in polynomial time \( x^1, \ldots, x^r \) satisfying the claimed constraints for \( k = 1, \ldots, r \). For \( r = 1 \), consider the system \( y^n \leq x^1 \leq y^1, \ Ax^1 = b \). We claim \( z := \frac{1}{n} \sum_{k=1}^n y^k \) is a real solution of this system. Since \( y^1, \ldots, y^n \) are feasible in (2), we have \( Az = \frac{1}{n} \sum_{k=1}^n Ay^k = b \). Since \( y^n \leq \cdots \leq y^1 \) as just proved, \( y^n \leq \frac{1}{n} \sum_{k=1}^n y^k \leq y^1 \). So \( z \) is a real solution. Since \( A \) is totally unimodular, this system also has an integer solution \( x^1 \), which can be found in polynomial time by linear programming. This proves the claim for \( r = 1 \).

Now we assume that the claim is valid for \( 1 \leq r < n \) and prove it for \( r + 1 \). By induction we can find in polynomial time \( x^1, \ldots, x^r \) satisfying the claimed constraints for \( k = 1, \ldots, r \). Consider the system in variable vector \( x^{r+1} \) with constraints

\[
y^n \leq x^{r+1} \leq y^1 , \ \sum_{k=n-r}^n y^k \leq \sum_{k=1}^{r+1} x^k \leq \sum_{k=1}^{r+1} y^k , \ Ax^{r+1} = b .
\]
We claim that \( z := \frac{1}{n-r} \left( \sum_{k=1}^{n} y^k - \sum_{k=1}^{r} x^k \right) \) is a real solution of this system.

First,

\[
Az = \frac{1}{n-r} \left( \sum_{k=1}^{n} Ay^k - \sum_{k=1}^{r} Ax^k \right) = \frac{1}{n-r} (nb - rb) = b.
\]

Second, by induction \( \sum_{k=1}^{r} x^k \leq \sum_{k=1}^{r} y^k \), and \( y^n \leq \cdots \leq y^1 \), and therefore

\[
z = \frac{1}{n-r} \left( \sum_{k=1}^{n} y^k - \sum_{k=1}^{r} x^k \right) \geq \frac{1}{n-r} \sum_{k=r+1}^{n} y^k \geq y^n.
\]

Third, by induction \( \sum_{k=1}^{r} x^k \geq \sum_{k=n+1-r}^{n} y^k \), and \( y^n \leq \cdots \leq y^1 \), and therefore

\[
z = \frac{1}{n-r} \left( \sum_{k=1}^{n} y^k - \sum_{k=1}^{r} x^k \right) \leq \frac{1}{n-r} \sum_{k=1}^{n-r} y^k \leq y^1.
\]

Fourth, \( \sum_{k=1}^{r} x^k \leq \sum_{k=1}^{r} y^k \) implies

\[
\frac{1}{n-r} \left( \sum_{k=1}^{r} y^k - \sum_{k=1}^{r} x^k \right) \leq \sum_{k=1}^{r} y^k - \sum_{k=1}^{r} x^k,
\]

and \( y^n \leq \cdots \leq y^1 \) implies \( \frac{1}{n-r} \sum_{k=r+1}^{n} y^k \leq y^{r+1} \). Therefore we obtain

\[
z = \frac{1}{n-r} \left( \sum_{k=1}^{r} y^k - \sum_{k=1}^{r} x^k + \sum_{k=n+1-r}^{n} y^k \right) \leq \left( \sum_{k=1}^{r+1} y^k - \sum_{k=1}^{r} x^k \right),
\]

which implies the inequality \( z + \sum_{k=1}^{r} x^k \leq \sum_{k=1}^{r+1} y^k \).

Fifth, \( \sum_{k=1}^{r} x^k \geq \sum_{k=n+1-r}^{n} y^k \) implies

\[
\frac{1}{n-r} \left( \sum_{k=1}^{n} y^k - \sum_{k=1}^{r} x^k \right) \geq \sum_{k=n+1-r}^{n} y^k - \sum_{k=1}^{r} x^k,
\]

and \( y^n \leq \cdots \leq y^1 \) implies \( \frac{1}{n-r} \sum_{k=1}^{n-r} y^k \geq y^{n-r} \). Therefore we obtain

\[
z = \frac{1}{n-r} \left( \sum_{k=n+1-r}^{n} y^k - \sum_{k=1}^{r} x^k + \sum_{k=1}^{n-r} y^k \right) \geq \left( \sum_{k=n-r}^{n} y^k - \sum_{k=1}^{r} x^k \right),
\]

which implies the inequality \( z + \sum_{k=1}^{r} x^k \geq \sum_{k=n-r}^{n} y^k \).

So \( z \) is a real solution. Since \( A \) is totally unimodular, this system also has an integer solution \( x^{r+1} \), which can be found in polynomial time by linear programming.
This completes the induction and the proof of Claim 2.

We now proceed with the proof of the theorem. First we solve the auxiliary program (2) and obtain optimal $y^1, \ldots, y^n$. This can be done in polynomial time using linear programming, since the matrix defining the equation system is $[A, \ldots, A]$ consisting of $n$ blocks of $A$, and is totally unimodular since $A$ is.

Next we find in polynomial time vectors $x^1, \ldots, x^n$ satisfying the conditions of Claim 2. Since $x^k$ is integer, $0 \leq y^n \leq x^k \leq y^1 \leq 1$, and $Ax^k = b$ for all $k$, the $x^k$ form a feasible solution in the program (1). Now we claim that $x^1 \land \cdots \land x^n = y^n$. Since $y^n \leq x^k$ for all $k$, we clearly have $x^1 \land \cdots \land x^n \geq y^n$. Now, the inequality $\sum_{i=n+1-k}^n y^i \leq \sum_{i=1}^k x^i \leq \sum_{i=1}^k y^i$ for $k = n$, which the $x^i$ and $y^i$ satisfy, gives $\sum_{i=1}^n x^i = \sum_{i=1}^n y^i$. But this implies that, for any coordinate $j$, if $x^1_j = \cdots = x^n_j = 1$ then also $y^1_j = \cdots = y^n_j = 1$, so if the $j$-th coordinate of $x^1 \land \cdots \land x^n$ is 1 then also $y^n_j = 1$. So also $x^1 \land \cdots \land x^n \leq y^n$, and we have equality $x^1 \land \cdots \land x^n = y^n$.

Now let $p := \sum_{k=1}^{n-1} k|y^k| + n^2 d|y^n|$ be the objective function value of the $y^k$ in program (2) and let $q := |x^1 \land \cdots \land x^n| = |y^n|$ be the objective function value of the $x^k$ in program (1). Then we have $p = \sum_{k=1}^{n-1} k|y^k| + n^2 d|y^n| \geq n^2 dq$. We now claim that $x^1, \ldots, x^n$ form an optimal solution for (1). Suppose indirectly $\bar{x}^1, \ldots, \bar{x}^n$ is a better solution with objective value $\bar{q} \leq q - 1$. By Claim 1, there is a feasible solution $\bar{y}^1, \ldots, \bar{y}^n$ in (2) with objective value $\bar{p}$ satisfying $\bar{p} < n^2 d(\bar{q}+1) \leq n^2 dq \leq p$ which contradicts the optimality of $y^1, \ldots, y^n$ in (2).

So summarizing, we can indeed find in polynomial time an optimal solution $x^1, \ldots, x^n$ for the intersection problem (1), completing the proof of the theorem.

Acknowledgments

The research of the first author was partially supported by a VPR Grant at the Technion and by the Fund for the Promotion of Research at the Technion. The research of the second author was partially supported by a Bourse d’Aide à la Mobilité Internationale from Université Paris Est.

References
